|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  26th Meeting: Geneva, CH, 12–20 January 2017 | Document: JCTVC-Z0025 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Spherical rotation orientation SEI for HEVC and AVC coding of 360 video** | | |
| *Status:* | Input Document | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Jill Boyce, Qian Xu | Tel: Email: | +1 503 712 2977 [jill.boyce@intel.com](mailto:jill.boyce@intel.com)  [qian.z.xu@intel.com](mailto:qian.z.xu@intel.com) |
| *Source:* | Intel | | |

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# Abstract

An SEI message is proposed for HEVC and AVC to indicate spherical rotation orientation of 360 degree video. As proposed, an encoder may perform spherical rotation of the input video prior to encoding, using up to 3 parameters (yaw, pitch, roll), in order to improve coding efficiency. The decoder can use the SEI message contents to perform the recommended inverse spherical rotation after decoding, before display. Up to -17.8% bitrate gain (using the WS-PSNR end-to-end metric) is shown for sequences in the JVET 360 video test conditions for HM16.14. The average for the entire test set is -2.9%, and many of the sequences do not benefit from the spherical rotation. The proposed syntax is independent of the particular projection format used, but the recommended spherical rotation operation relies on having knowledge of the projection format. In JVET-E0075, the same proposal is made for the JEM, but also includes an option to include the orientation parameters in the PPS, instead of in an SEI.

# Introduction

360 degree video can be coded using existing HEVC and AVC profiles, using a projection format to map spherical video data to a rectangular format used for coding. The projection mapping operations treat different regions of an input sphere differently, and introduce warping and artificial discontinuities. The block motion prediction used in HEVC and AVC were not specifically designed to code warped and discontinuous regions efficiently.

For equi-rectangular projection (ERP), image areas near the north and south poles are warped, and an artificial discontinuity is introduced at the left and right edges. Objects in the video that are projection mapped around the north and south pole areas, or split across the left and right edges, will not code as efficiently as objects that are in the center of the ERP frame, because they will be subject to less warping and avoid an artificial discontinuity.

Other projection formats such as Cube Map Projection (CMP) introduce artificial discontinuities at cube face edges. CMP also introduces warping near the cube face edges, but to a lesser degree than the warping found in ERP.

Experimental results have shown that coding efficiency of 360 degree video can be improved for some test sequences by applying spherical rotation prior to encoding. An SEI message is proposed, in which 3 spherical rotation parameters (yaw, pitch, and roll) may be signaled, for a decoder to apply after decoding and prior to display, after the projection mapping has been applied to map the decoded video picture to a sphere. This operation can be combined with the viewport generation. The proposed syntax and semantics are described in section 2.

An encoder algorithm was developed to automatically determine if spherical rotation should be applied, and if so, to select the (yaw, pitch, roll) parameters applied to the input video prior to encoding. The algorithm is described in section 3. Software is provided with the contribution.

Experimental results for the JVET-D1030 conditions are provided, using HM16.14, in Section 4. Results for the JEM are also provided in JVET-E0075.

Figure 1 shows the first frame of the Chairlift sequence. Figure 2 shows that frame after a rotation of (79, 12, 0) degrees has been applied.

|  |
| --- |
| ChairLift_beforeRotation |
| **Figure 1. First frame of Chairlift sequence** |

|  |
| --- |
| ChairLift_afterRotation |
| **Figure 2. First frame of Chairlift sequence, after rotation of (79, 12, 0) degrees** |

|  |
| --- |
| Image result for pitch yaw roll xyz |
| **Figure 3 Yaw, pitch, and roll rotations** |

HEVC and AVC already include a display orientation SEI message, for use with regular 2-D video, which provides to the decoder a recommended 2-D rotation and/or flip to be applied after decoding and before display.

Although a standalone SEI message is proposed, an alternative is to add the spherical orientation parameters to the planned omnidirectional\_projection SEI.

In JVET-E0075, spherical rotation is proposed as either an SEI message or a normative operation with parameters carried in the picture parameter set.

# Proposed syntax and semantics

The proposed syntax and semantics describe the recommended spherical rotation to be applied at the decoder. The SEI message may be changed for each coded picture, but can persist until it is cancelled or replaced.

The proposed syntax is independent of the particular projection format used, but the recommended spherical rotation operation relies on having knowledge of the projection format.

## Syntax

|  |  |
| --- | --- |
| spherical\_orientation( payloadSize ) { | **Descriptor** |
| **spherical\_orientation\_cancel\_flag** | u(1) |
| if( !spherical\_orientation\_cancel\_flag ) { |  |
| **spherical\_orientation\_precision** | ue(v) |
| **spherical\_orientation\_yaw** | se(v) |
| **spherical\_orientation\_pitch** | se(v) |
| **spherical\_orientation\_roll** | se(v) |
| **spherical\_orientation\_persistence\_flag** | u(1) |
| } |  |
| } |  |

## Semantics

When the associated picture has PicOutputFlag equal to 1, the spherical orientation SEI message informs the decoder of a transformation that is recommended to be applied to a spherical mapping of the cropped decoded picture prior to display. If the omnidirectional\_projectionSEI message is present in an access unit, the spherical\_orientation SEI should follow it. The recommended projection mapping indicated by the omnidirectional\_projectionSEI should be performed first to map the cropped decoded picture to a sphere, and then the recommended spherical rotation should be applied.

**spherical\_orientation\_cancel\_flag** equal to 1 indicates that the SEI message cancels the persistence of any previous spherical orientation SEI message in output order. spherical\_orientation\_cancel\_flag equal to 0 indicates that spherical orientation information follows.

**spherical\_orientation\_precision** specifies the precision of spherical\_orientation\_pitch, spherical\_orientation\_yaw, and spherical\_orientation\_roll. spherical\_orientation\_precision shall be in the range of 0 to 15 inclusive.

**spherical\_orientation\_yaw** specifies the recommended spherical rotation of the projection mapped decoded picture along the Z-axis prior to display. The projection mapped decoded picture should be rotated along the Z-axis by 180\* spherical\_orientation\_yaw ÷ 2spherical\_orientation\_precision degrees (π \* spherical\_orientation\_yaw  ÷ 2sperhical\_orientation\_precision radians, where π is Archimedes' constant 3.141 592 653 589 793...) in the anticlockwise direction prior to display.

**spherical\_orientation\_pitch** specifies the recommended spherical rotation of the projection mapped decoded picture along the Y-axis prior to display. The projection mapped decoded picture should be rotated along the Y-axis by 180 \* spherical\_orientation\_pitch  ÷ 2spherical\_orientation\_precision degrees (π \* spherical\_orientation\_pitch ÷ 2sperhical\_orientation\_precision radians, in the anticlockwise direction prior to display.

**spherical\_orientation\_roll** specifies the recommended spherical rotation of the projection mapped decoded picture along the X-axis prior to display. The projection mapped decoded picture should be rotated along the X-axis by 180 \* spherical\_orientation\_roll  ÷ 2spherical\_orientation\_precision degrees (π \* spherical\_orientation\_roll  ÷ 2sperhical\_orientation\_precision radians, in the anticlockwise direction prior to display.

**spherical\_orientation\_persistence\_flag** specifies the persistence of the spherical orientation SEI message for the current layer.

spherical\_orientation\_persistence\_flag equal to 0 specifies that the spherical orientation SEI message applies to the current decoded picture only.

Let picA be the current picture. spherical\_orientation\_persistence\_flag equal to 1 specifies that the spherical orientation SEI message persists for the current layer in output order until one or more of the following conditions are true:

– A new CLVS of the current layer begins.

– The bitstream ends.

– A picture picB in the current layer in an access unit containing a spherical orientation SEI message that is applicable to the current layer is output for which PicOrderCnt( picB ) is greater than PicOrderCnt( picA ), where PicOrderCnt( picB ) and PicOrderCnt( picA ) are the PicOrderCntVal values of picB and picA, respectively, immediately after the invocation of the decoding process for picture order count for picB.

# Encoder algorithm

An automatic algorithm was developed to determine if spherical rotation should be applied to a video sequence. If spherical rotation was selected, (pitch, yaw, roll) parameters were selected to attempt to improve the coding efficiency of the sequence. Software is included with the contribution. Encoder pre-processing is performed for the first GOP to select the parameters, which are then applied to the entire sequence. All non-intra frames in the first GOP are included to derive the motion information.

Denote the longitude as ϕ in the range [-π, π], and the latitude as θ in the range [-π/2, π/2]. The algorithm checks the amount of motion in the video at (ϕ, θ) and its symmetric point (,) around the center of the sphere. θ= π/2 and correspond to the current north and south poles respectively. The algorithm searches for the pair of areas (ϕ, θ) and (,) with the smallest motion and rotates them to the north and south poles respectively.

To find such a pair, we first divide the ERP frame into multiple 4x4 CTUs, each centered on a sampling position of (m, n) (0≤ m <W, 0 ≤ n < H). The mapping between a sampling position (m, n) in the ERP and the (ϕ, θ) can be found in [D0090]. We calculate the amount of motion for each CTU and then sum up the 4x4 CTUs. Specifically, for each prediction unit of CTU(j, i), the amount of motion is calculated as

.

For an intra PU, MVx and MVy are set to 200.

The motion for CTU(j, i) is then

A similar process is performed to calculate the motion for sampling positions (, which corresponds to (,) in the unit sphere.

After we select the sampling position ( with the minimum motion, it is compared to the motion of the current north and south poles. The motion is classified into low, medium, and high levels based on empirical values. If the current motion is at least medium level and the minimum motion found is at low level, it is determined to rotate ( to the north pole, otherwise it is better to keep the current sphere.

# Experiments

The JVET common test conditions and evaluation procedures for 360° video described in JVET-D1030 were used in the experiments, using HM16.14, for ERP format.

For the test set, spherical rotation was selected for only the Chairlift and DrivingInCountry sequences, which showed gains of -17.8% and -10.8% respectively for the WS-PSNR E2E metric. Averaging for the 10 sequences, an average gain of -2.9% is shown.

These results were computed before consideration of the SEI message syntax bits, although the provided software includes the SEI message syntax insertion. For the two sequences where the SEI message was used, the added bits for the SEI message were 40 bits.

|  |  |  |  |
| --- | --- | --- | --- |
| ERP rotation vs. ERP | WS-PSNR (End to End) | | |
| Y | U | V |
| Train | 0.0% | 0.0% | 0.0% |
| Skateboarding\_trick | 0.0% | 0.0% | 0.0% |
| Skateboarding\_in\_lot | 0.0% | 0.0% | 0.0% |
| Chairlift | -17.8% | -8.8% | -10.3% |
| KiteFlite | 0.0% | 0.0% | 0.0% |
| Harbor | 0.0% | 0.0% | 0.0% |
| PoleVault | 0.0% | 0.0% | 0.0% |
| AerialCity | 0.0% | 0.0% | 0.0% |
| DrivingInCity | 0.0% | 0.0% | 0.0% |
| DrivingInCountry | -10.8% | -15.3% | -15.9% |
| Overall | -2.9% | -2.4% | -2.6% |

The HM16.14 results for the additional metrics are below.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ERP rotation  vs. ERP | S-PSNR-NN | | | WS-PSNR | | | S-PSNR-I | | | CPP-PSNR | | |
| Y | U | V | Y | U | V | Y | U | V | Y | U | V |
| Train | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| Skateboarding  \_trick | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| Skateboarding  \_in\_lot | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| Chairlift | -17.2% | -9.1% | -10.1% | -17.3% | -9.0% | -10.5% | -16.3% | -9.4% | -9.5% | -16.4% | -9.3% | -10.0% |
| KiteFlite | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| Harbor | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| PoleVault | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| AerialCity | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| DrivingInCity | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% | 0.0% |
| DrivingInCountry | -11.1% | -15.2% | -15.5% | -11.3% | -15.2% | -15.7% | -11.0% | -14.8% | -15.0% | -11.3% | -14.7% | -15.0% |
| Overall | -2.8% | -2.4% | -2.6% | -2.9% | -2.4% | -2.6% | -2.7% | -2.4% | -2.4% | -2.8% | -2.4% | -2.5% |

# Patent rights declaration(s)

**Intel may have current or pending patent rights relating to the technology described in this contribution and, conditioned on reciprocity, is prepared to grant licenses under reasonable and non-discriminatory terms as necessary for implementation of the resulting ITU-T Recommendation | ISO/IEC International Standard (per box 2 of the ITU-T/ITU-R/ISO/IEC patent statement and licensing declaration form).**